**Instructions:**
Be verbose. Explain clearly your reasoning, methods, and results in your written work. Write clear code that is well documented. With 99% certainty, you cannot write too many code comments.

Written answers are worth 8 points. Code is worth 2 points. 10 points total.

1. When finished, respond to the question in Canvas as "done." We will record your grade there.
2. In your code repository, create a folder called "Week02."
3. In that folder, include
    a. a document (PDF) with your responses.
    b. All code
    c. A README file with instructions for us to run your code

Everything must be checked into your repository by 8am Sunday 9/14. A pull will be done at that time. Documents and code checked in after the instructors pull will not be graded.

Data for problems can be found in CSV files with this document in the class repository.

**Problem 1 (3pts)**
Given the dataset in problem1.csv

1. Calculate the first 4 moment values using the normalized formulas in the Week 1 notes.
2. Calculate the first 4 moment values using your chosen statistical package.
3. Are your statistical package functions biased? Prove or disprove your hypotheses. Explain your conclusion.

**Problem 2 (5pts)**
Assume the multiple linear regression model $Y = X\boldsymbol{\beta} + \epsilon$

1. Fit the data in problem2.csv using OLS. Then fit the data using MLE given the assumption of normality. Compare the beta values and the standard deviation of the OLS errors to the fitted MLE $\sigma$. What is your finding? Explain any differences.
2. Fit the data in problem2.csv using MLE given the assumption of a T distribution of errors. Show the fitted parameters. Compare the fitted parameters among the MLE under the normality assumption and T distribution assumption. Which is the best fit?
3. Fit a multivariate distribution to the data in problem2_x.csv. Given the values of $X_1$ what are the conditional distributions for $X_2$ for each observation. Plot the expected value along with the 95% confidence interval and the observed value.
4. (1 point Extra Credit). $Y = X\boldsymbol{\beta} + \epsilon \; and \; \epsilon \sim N(0, \sigma^2)$. Derive the maximum likelihood estimators for $\beta$ and $\sigma^2$

**Problem 3 (2pts)**

Examine the data in problem3.csv; which AR(n) or MA(n) model do you expect to fit this data best? Fit the data using AR(1) - AR(3) and MA(1) - MA(3) models. Which is the best fit and does this confirm your hypothesis?